

UNIVERSIDAD TECNOLÓGICA DE PEREIRA
FACULTAD DE INGENIERÍAS

MAESTRÍA EN INGENIERÍA DE SISTEMAS Y COMPUTACIÓN

ANÁLISIS MULTIVARIADO

OBJETIVO GENERAL

El curso es de un nivel matemático intermedio y tiene una orientación teórico- práctica. Con esto se pretende que el estudiante desarrolle habilidades intuitiva y analíticas para así resolver problemas estadísticos en el campo multivariado, conozca la estructura básica de una demostración y se ejercite en la formulación de conclusiones razonables a partir de los resultados obtenidos con las diversas técnicas estudiadas. Durante el curso el estudiante aprenderá a usar procedimientos multivariados implementados en paquetes estadísticos tales como SAS® y R y a demostrar algunos de los resultados clásicos de la teoría del análisis multivariado; además, adquirirá herramientas básicas para realizar investigación estadística.

Estudiar datos multidimensionales desde el punto de vista computacional con grandes bases de datos, chequeos distribucionales que incluyen gran cantidad de teoría y métodos estadísticos, tales como: Algoritmos Genéticos, Bootstrap, Métodos Montecarlo y Cadenas de Markov (MCMC), Análisis Bayesiano, Modelos Lineales Generalizados, Verosimilitud e Inferencia Estadística.

CONTENIDO

MODULO I. INTRODUCCIÓN

- 1.1 Motivación
- 1.2 Definiciones
- 1.3 Repaso de álgebra lineal, Teorema de la descomposición espectral

MODULO II. OPERADORES VALOR ESPERADO Y COVARIANZA

- 2.1 Demostración de algunas propiedades
Del valor esperado y la Covarianza
- 2.2 Definición de dispersión, matriz de correlación y sus propiedades
- 2.3 Dispersión particionada

MODULO III. ORGANIZACIÓN Y PRESENTACIÓN DE DATOS MULTIVARIADOS

- 3.1 Definiciones
- 3.2 Estadísticos descriptivos

3.3 Algunas representaciones gráficas de datos

MODULO IV. DISTRIBUCIÓN NORMAL MULTIVARIADA

- 4.1 Definición
- 4.2 Geometría de la normal multivariada
- 4.3 Resultados y propiedades relacionadas con la normal multivariada
- 4.4 Función generadora de momentos de la normal multivariada
- 4.5 Evaluando el supuesto de normalidad
- 4.6 Detección de outliers y transformaciones para alcanzar normalidad
- 4.7 Distribución condicional normal multivariada
- 4.8 Generación aleatoria de normales multivariadas

MODULO V. DISTRIBUCIONES MUESTRALES

- 5.1 Definición de muestra aleatoria multivariada
- 5.2 Propiedades del vector de medias y la matriz de varianzas y Covarianzas muestrales
- 5.3 Definición de sesgo multivariado
- 5.4 Definición de error cuadrático medio y propiedades. Ejemplos
- 5.5 Consistencia de un estimador multivariado
- 5.6 Datos centrados. Ejemplos

MODULO VI. ESTIMACIÓN DE MÁXIMA VEROSIMILITUD

- 6.1 Definiciones
- 6.2 Estimadores de máxima verosimilitud del vector de medias y la matriz de varianzas y Covarianzas muestrales
- 6.3 Función score y función de información

MODULO VII. DISTANCIA DE MAHALANOBIS

- 7.1 Definición de distancia estándar
- 7.2 Propiedades de la distancia de Mahalanobis
- 7.3 Teorema fundamental de la distancia de Mahalanobis

MODULO VIII. ANÁLISIS DE COMPONENTES PRINCIPALES Y CLUSTERING (AGRUPAMIENTO)

- 8.1 Definición
- 8.2 Propiedades de la reducción de dimensionalidad
- 8.3 Componentes principales muestrales
- 8.4 Inferencia para las componentes muestrales. Ejemplos
- 8.5 Métodos jerárquicos: Vínculo simple, Vínculo completo, Vínculo promedio
- 8.6 Caras de Chernoff; estrellas

MODULO IX. INFERENCIA ESTADÍSTICA PARA LAS MEDIAS

- 9.1 TEST DE HIPÓTESIS PARA MUESTRAS PEQUEÑAS

- 9.2 Estadístico T2 de Hotelling, propiedades
- 9.3 Relación entre el estadístico F y la T2 de Hotelling
- 9.4 Reglas para probar hipótesis acerca de un vector de medias en el caso de una muestra aleatoria de vectores
- 9.5. Pruebas de razón de verosimilitud: el estadístico lambda de Wilks
- 9.6 Equivalencia entre el estadístico T2 de Hotelling y el lambda de Wilks
- 9.7. Distribución del estadístico lambda de Wilks
- 9.8. Inferencias acerca del vector de medias poblacionales con muestras grandes
- 9.9. Pruebas de hipótesis para comparar los vectores de medias en dos poblaciones:
dispersiones iguales y dispersiones diferentes
- 9.10 Inferencia para la matriz de dispersión

MODULO X. INTRODUCCIÓN AL ANÁLISIS DISCRIMINANTE

- 10.1 Separación y clasificación para dos poblaciones
- 10.2 Función lineal discriminante de Fisher
- 10.3 Regresión logística

COMPETENCIAS

COMPETENCIAS TRANSVERSALES / GENÉRICAS:

- Aprendizaje autónomo
- Capacidad de análisis y síntesis
- Capacidad de aplicar los conocimientos a la práctica
- Resolución de problemas
- Trabajo individual y por parejas
- Comunicación oral y escrita

COMPETENCIAS ESPECÍFICAS:

- Cognitivas (Saber):
 - Idioma
 - Matemáticas
 - Nuevas tecnologías TIC
 - Conocimientos de informática
- Procedimentales / Instrumentales (Saber hacer):
 - Redacción en interpretación de documentación técnica
 - Estimación y programación del trabajo
 - Planificación, organización y estrategia.
- Actitudinales (Ser):
 - Calidad
 - Toma de decisión
 - Capacidad de iniciativa y participación

TÉCNICAS DOCENTES

Las técnicas docentes que se van a utilizar son:

- Clases de teoría
- Exposiciones sobre trabajos de casos prácticos.
- Tutorías colectivas de teoría
- Clases de prácticas
- Corrección de las prácticas
- Tutorías colectivas de prácticas
- Tutorías individualizadas

DESARROLLO Y JUSTIFICACIÓN:

Clases de teoría:

- Se hará una reseña inicial del contenido de cada tema y se indicará su relación con los otros temas.
- Al comenzar la explicación de una sección de un tema, se indicarán las relaciones que posee con otras secciones del mismo tema o de temas diferentes.
- Se explicará detenidamente cada sección de cada tema teórico.

Exposiciones:

- El profesor propondrá los trabajos sobre trabajos de casos prácticos, que los estudiantes deberán preparar y exponer a lo largo del curso.
- Los trabajos podrán hacerse individualmente o en parejas.

Acerca de las prácticas:

Las prácticas y tutorías se realizarán por grupos en el laboratorio con base en la guía del curso.

Tutorías colectivas de teoría o prácticas

Es una actividad desarrollada dentro de las horas de clase

- El profesor responderá a las preguntas que les planteen los estudiantes procurando que ellos intenten deducir las repuestas correctas.
- Se procurará que las preguntas que se planteen no sean dudas particulares de un estudiante, sino dudas generales que puedan tener la mayoría de los estudiantes. Las dudas particulares se deben plantear en las tutorías individuales.

- El profesor también podrá plantear preguntas a los estudiantes para comprobar si han aprendido correctamente los conceptos fundamentales de la asignatura.

Tutorías individualizadas:

Según es reglamento estudiantil vigente, en su artículo 60. (“ARTÍCULO 60o.: El estudiante de la Universidad tiene derecho a:.....Ser asistido, asesorado y oído por quienes tienen la responsabilidad administrativa y docente.”. Subrayado nuestro), estas tutorías están enmarcadas dentro de la actividad docente y los horarios deberán ser concertados con todos los estudiantes o con la mayoría cuando con todos no sea posible.

- Los estudiantes con el fin de poder organizar y garantizar que la atención sea individual, deberá solicitar con anticipación cita con el profesor.
- Los estudiantes deben utilizar estas tutorías a lo largo de todo el curso y no sólo antes de la fecha del examen.
- El profesor intentará resolver las dudas particulares que pueda tener cada estudiante en relación con los temas de teoría, los trabajos de las exposiciones, las prácticas, etc.
- Aunque las dudas más simples puedan plantearse mediante correo electrónico, es preferible que haya una reunión del profesor y el estudiante para resolver las dudas más complejas.
- La Universidad podrá disponer como recurso adicional un “asistente de cátedra o monitor”, que podrá ser un estudiante de semestres superiores, según el reglamento que sobre este particular maneje la Universidad.

MECANISMOS DE CONTROL Y SEGUIMIENTO

El profesor podrá comprobar el grado de seguimiento de la asignatura mediante:

- La asistencia a las clases de teoría y prácticas
- Las exposiciones de temas de teoría.
- La corrección de las prácticas.
- Las tutorías personales
- Los parciales
- Los exámenes de corta duración (Quiz).

EVALUACIÓN

Taller 1 ó Examen Corto 1

20%

Taller 2 ó Examen Corto 2	20%
Trabajo Final	30%
Examen Final	30%.

Los porcentajes están sujetos a ser negociados con los alumnos por parte del docente, exceptuando el porcentaje del examen final que nunca deberá ser inferior al 30%.

BIBLIOGRAFÍA

Johnson, R. And Wichern, D.W. Applied Multivariate Statistical Analysis. Fourth edition. Prentice Hall Inc., 1988
 Seber, G.A.F. Multivariate Observations. John Wiley & Sons. 1984
 Flury, B. A first course in multivariate statistics. Springer texts in statistics. 1997